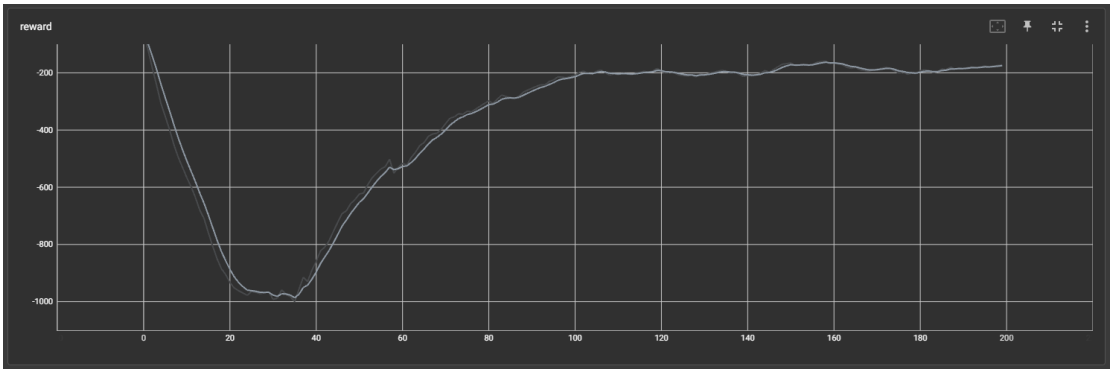
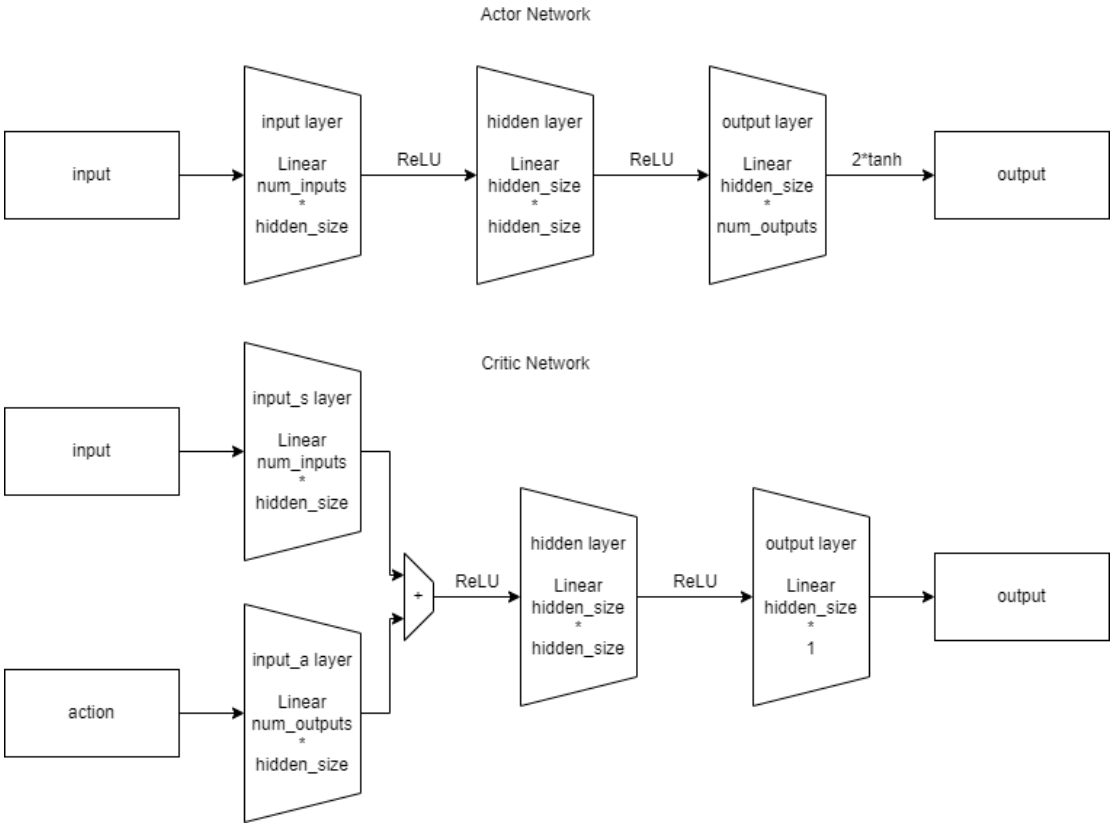
Hw2 report

1. Pendulum-v1
   a、 Results: rewards around -100~-200



   b、 Hyperparameters:
      i. Episodes=200
      ii. Gamma=0.995
      iii. Tau=0.002
      iv. Hidden size=128
      v. Noise scale=0.3
      vi. Batch size=128
      vii. Actor learning rate=1e-4
      viii. Critic learning rate=1e-3
      ix. NN structure:

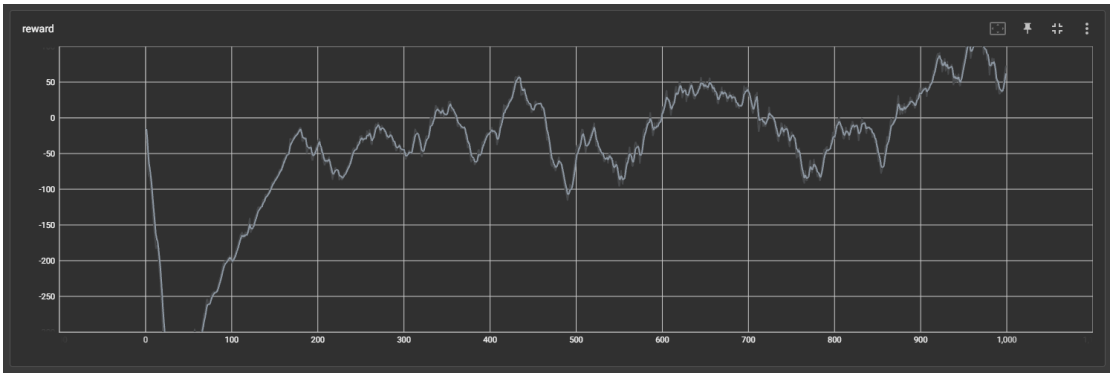


c、 Experiments

i. I tried several types of networks, including dropout layer, multiple hidden layers, but their performance didn't stand out much. So, I pick the simplest structure and add some ReLU, then it performs quite well.

2. LunarLanderContinuous-v2

a、 Result: rewards around 100



i. Episodes=1000

ii. Gamma=0.995

iii. Tau=0.001

iv. Hidden size=300

v. Noise scale=1.2

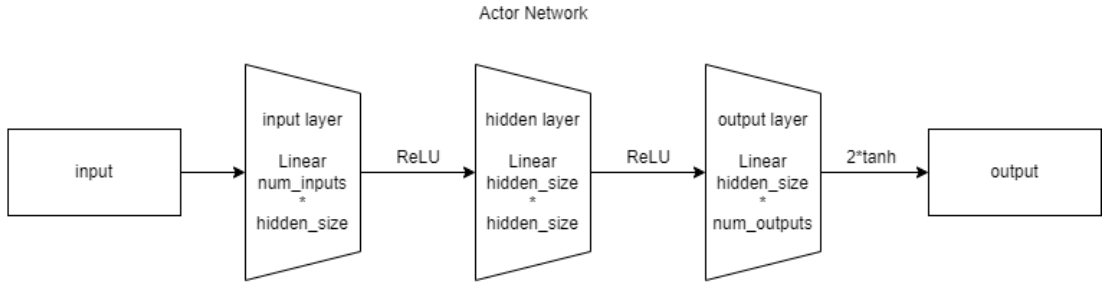
vi. Batch size=256

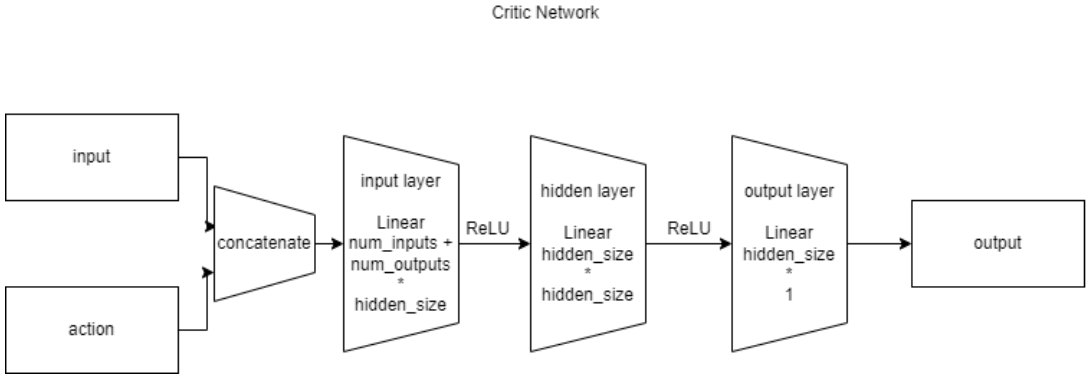
vii. Actor learning rate=3e-4

viii. Critic learning rate=1e-4

ix. Epsilon=0.15

x. NN structure:





b、 Experiments

i. Like above, I tried some compound networks, but both of them performs badly, so I decided to use the simplest structure and tune the parameters.

ii. I discovered that the lander always floating above the ground and not landing, so I decided to use epsilon-greedy to choose the action. After that, the lander landed quite well.